The Pigeonhole Principle (PP)

(Simple form) Let n be any positive integer. If at least $n + 1$ objects are distributed among n boxes, then (at least) one of the boxes must contain at least 2 objects.
(Strong form) Let k and n be any two positive integers. If at least $kn + 1$ objects are distributed among n boxes, then (at least) one of the boxes must contain at least $k + 1$ objects.

<u>Proof</u>

If no boxes contain $k + 1$ or more objects, then every box contains at most k objects. This implies that the total number of objects put into the n boxes is at most kn, a contradiction.

Alternatively, we can also formulate the pigeonhole principle as:

If m objects are distributed among n boxes, then (at least) one of the boxes must contain at least $\left\lceil \frac{m}{n} \right\rceil$ objects.

Two main skills are required. Firstly we need to work through the logic of the solution, which often includes identifying appropriate "pigeons" and "pigeonholes". Secondly we need to write the proof in a form that is readable and rigourous (and not too long). The two are somewhat connected; often a clear definition of appropriate variables or objects helps with both.

Example 1
Show that among any 4 integers, there are two whose difference is divisible by 3.

Solution

Let the four integers be a, b, c, d.

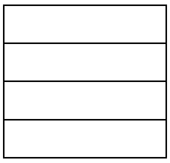
Since each must be $\equiv 0, 1, \text{or } 2 \pmod 3$, by pigeonhole principle there must be two of them that are congruent to the same remainder mod 3; without loss of generality let these be a and b.

Then $a \equiv b \pmod 3$ (by transitivity of $\equiv$), so $a - b \equiv 0 \pmod 3$ and $a - b$ is divisible by 3. □

Example 2
Show that for any set of 5 points chosen within a square whose sides are of length 2 units, there are two points in the set whose distance apart is at most $\sqrt{2}$.

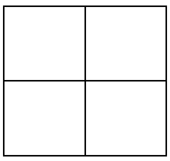
Thought process

With 5 points, we will be able to partition the square into 4 pieces and use pigeonhole principle to conclude "there must be at least 2 points in one of the partitions". So, for the proof to work, we must create partitions such that the distance between any 2 points in the same partition is not more than $\sqrt{2}$. In particular, the below is not helpful:



Solution

Divide the square into 4 sub-squares of unit length:



Since there are 5 points, by pigeonhole principle there must be one sub-square with at least 2 points in/on it. But the distance between any two points in/on the same sub-square is at most $\sqrt{2}$. So, the two points in/on the same sub-square are distance at most $\sqrt{2}$ apart. □

Example 3

Let $B = \{b_1, b_2, b_3, b_4, b_5\}$ be a set of 5 distinct positive integers. Show that for **any** permutation (i.e. rearrangement of the elements) $b_{i_1} b_{i_2} b_{i_3} b_{i_4} b_{i_5}$ of *B*, the product

$$(b_{i_1} - b_1)(b_{i_2} - b_2) \ldots (b_{i_5} - b_5)$$

is always even.

For example, if *B* = {26, 7, 1, 22, 11} and a permutation of *B* is (1, 22, 26, 11, 7). Then the product (1 – 26)(22 – 7)(26 – 1)(11 – 22)(7 – 11) is even.

Thought process

A product is even if and only if one or more factors is even. The fact that 5 is odd seems to be important, since that means there must be more of one type…

Solution 1

By pigeonhole principle, at least $\left\lceil \frac{5}{2} \right\rceil = 3$ of the integers must be of the same parity. WLOG suppose b_1, b_2, b_3 are of the same parity.

Then a maximum of two of b_{i_4}, b_{i_5} can be b_1, b_2 or b_3. That implies that one of $b_{i_1}, b_{i_2}, b_{i_3}$ is equal to b_1, b_2 or b_3. So one of $\left(b_{i_1} - b_1\right), \left(b_{i_2} - b_2\right), \left(b_{i_3} - b_3\right)$ is a difference of two integers of the same parity, which means the difference is an even integer. So the entire product is even. □

Thought process

While the above is a valid proof, the second half of the solution feels a little awkward to write. In this particular case, the somewhat unnatural step of combining the permuted and non-permuted integers makes the solution shorter.

Solution 2

By pigeonhole principle, at least $\left\lceil \frac{5}{2} \right\rceil = 3$ members of *B* have the same parity. This means that at least 6 of $b_1, b_2, b_3, b_4, b_5, b_{i_1}, b_{i_2}, b_{i_3}, b_{i_4}, b_{i_5}$ have the same parity.

By pigeonhole principle, when these 10 integers are distributed into 5 factors, there must be at least two of them in the same factor, say $b_{i_k} - b_k$. Then this factor is even, so the entire product is even. □

Example 4 (Handshake Problem, also appeared as 2018/Q6a)
Consider any group of n students where $n \geq 2$. Prove that there are at least two students who have the same number of friends in the group. State clearly any assumption you have made about friendships.

Assumptions
We assume that friendship is:
- non-reflexive (no one is a friend of themself) and
- symmetric (if A is a friend of B, then B is a friend of A).

Thought process

With the two assumptions, the number of friends a person can have ranges from 0 to $n-1$. At first sight, this is not enough because we have n people and n possible number of friends. By thinking deeper about the "0 friends" case, we might find a breakthrough.

Initial Solution

Case 1: There is at least 1 student with no friends. Then the maximum number of friends a student can have is $n-2$. So there are only $n-1$ different possible numbers of friends, namely $\{0, 1, ..., n-2\}$, among n students. By pigeonhole principle, there must be two students who have the same number of friends.
Case 2: There are 0 students with no friends. Then the minimum number of friends a student can have is 1. So there are only $n-1$ different possible numbers of friends, namely $\{1, 2, ..., n-1\}$, among n students. By pigeonhole principle, there must be two students who have the same number of friends. □

Thought process

Looking back at the cases, the crux is that there cannot be both a student with 0 friends and a student with $n-1$ friends. But this means that we can lump both into one pigeonhole.

Refined Solution

Divide the students into $n-1$ categories A_k, $k = 1, 2, ..., n-1$ as follows:

A_1	A_2	...	A_{n-2}	A_{n-1}
1 friend	2 friends	...	$n-2$ friends	0 or $n-1$ friends

By pigeonhole principle, there must be at least 2 students in 1 category. If there are 2 students in any of the first $n-2$ categories, they must have the same number of friends. For the last category, if there are 2 students in that category, they must also have the same number of friends because it is not possible for one student to have $n-1$ friends (so is friends with everyone) and another to have 0 friends at the same time. □

One lesson from Example 4 is that sometimes when we don't seem to have enough pigeons, we can think about looking for more pigeons or combining some pigeonholes.

Example 5 (2018/Q6b)
Let x be any positive real number and let n be any positive integer. Prove that there are integers a and b, with $1 \le b \le n$, such that $\left| x - \frac{a}{b} \right| < \frac{1}{bn}$.

Thought process

The condition $\left| x - \frac{a}{b} \right| < \frac{1}{bn}$ can be re-written as $|bx - a| < \frac{1}{n}$. Since a is integer, this hints at examining the fractional part of bx as b varies.
A natural thing is to divide the unit interval into segments of length $\frac{1}{n}$. But there are n intervals, and we only have n different values of bx. Keeping this unresolved issue in mind, let us first confirm that the logic works if any pigeonhole contains two pigeons:

Trial Solution

For any positive number y, we define $\operatorname{frac}(y) = y - \lfloor y \rfloor$.
Consider the n numbers $\operatorname{frac}(x), \operatorname{frac}(2x), \ldots, \operatorname{frac}(nx)$ and the n intervals

$$\left[0, \frac{1}{n}\right), \left[\frac{1}{n}, \frac{2}{n}\right), \left[\frac{2}{n}, \frac{3}{n}\right), \ldots, \left[\frac{n-1}{n}, 1\right).$$

Suppose one interval contains both $\operatorname{frac}(px)$ and $\operatorname{frac}(qx)$ where $p > q$.

Since the difference between two numbers in the same interval must be smaller than $\frac{1}{n}$,

$$\left| \operatorname{frac}(px) - \operatorname{frac}(qx) \right| < \frac{1}{n}$$

$$\left| px - \lfloor px \rfloor - qx + \lfloor qx \rfloor \right| < \frac{1}{n}$$

$$\left| (p-q)x - \left(\lfloor px \rfloor - \lfloor qx \rfloor \right) \right| < \frac{1}{n}$$

$$\left| (p-q)\left(x - \frac{\lfloor px \rfloor - \lfloor qx \rfloor}{p-q} \right) \right| < \frac{1}{n}$$

$$\left| x - \frac{\lfloor px \rfloor - \lfloor qx \rfloor}{p-q} \right| < \frac{1}{(p-q)n}$$

So, letting $a = \lfloor px \rfloor - \lfloor qx \rfloor$ and $b = p - q$, we get the required result $\left| x - \frac{a}{b} \right| < \frac{1}{bn}$.

Thought process

So the general case works, and now we need to tackle the gap. After some thought, we might realise that the first interval works a little bit differently; we don't need two pigeons in this interval to construct our *a* and *b*, just one:

If $\text{frac}(kx) < \frac{1}{n}$ then

$$0 \le kx - \lfloor kx \rfloor < \frac{1}{n}$$

$$0 \le x - \frac{\lfloor kx \rfloor}{k} < \frac{1}{kn}$$

so we can simply take $a = \lfloor kx \rfloor$ and $b = k$.

Now we can divide into two cases: either one number falls within the first interval and we use the above construction, or no number falls within the first interval and we only have $n-1$ intervals so pigeonhole principle applies.

But: before rushing into writing the proof, do we notice how similar the steps for the first interval are to the general case? Is there some way to combine and simplify our answer?

Final Solution

For any positive number *y*, we define $\text{frac}(y) = y - \lfloor y \rfloor$.

Consider the $n+1$ numbers $\text{frac}(0), \text{frac}(x), \text{frac}(2x), \ldots, \text{frac}(nx)$ and the *n* intervals

$$\left[0, \frac{1}{n}\right), \left[\frac{1}{n}, \frac{2}{n}\right), \left[\frac{2}{n}, \frac{3}{n}\right), \ldots, \left[\frac{n-1}{n}, 1\right).$$

By pigeonhole principle, there must be 2 numbers that lie in the same interval. WLOG let these numbers be $\text{frac}(px)$ and $\text{frac}(qx)$ where $p > q$.

Since the difference between two numbers in the same interval must be smaller than $\frac{1}{n}$,

$$\left|\text{frac}(px) - \text{frac}(qx)\right| < \frac{1}{n}$$

$$\left|px - \lfloor px \rfloor - qx + \lfloor qx \rfloor\right| < \frac{1}{n}$$

$$\left|(p-q)x - (\lfloor px \rfloor - \lfloor qx \rfloor)\right| < \frac{1}{n}$$

$$\left|(p-q)\left(x - \frac{\lfloor px \rfloor - \lfloor qx \rfloor}{p-q}\right)\right| < \frac{1}{n}$$

$$\left|x - \frac{\lfloor px \rfloor - \lfloor qx \rfloor}{p-q}\right| < \frac{1}{(p-q)n}$$

So, letting $a = \lfloor px \rfloor - \lfloor qx \rfloor$ and $b = p - q$, we get the required result $\left|x - \frac{a}{b}\right| < \frac{1}{bn}$. □

Example 6
Show that, when a 9×15 checkerboard is completely tiled by non-overlapping 1×3 dominos, at least one line between two adjacent rows or two adjacent columns is covered by at least 5 dominos.

Thought process

The number of tilings is very large; dividing into cases is unlikely to be the way to go. Instead, we would like to think about "global" arguments, which pigeonhole principle is well-suited for. But whether this works depends entirely on the size of the numbers involved.

Solution

The number of dominos is $\dfrac{9 \times 15}{1 \times 3} = 45$, and each domino covers 2 line segments for a total of 90 line segments covered.
There are 8 lines between adjacent rows and 14 lines between adjacent columns, for a total of 22 lines.
Since $\left\lceil \dfrac{90}{22} \right\rceil = \lceil 4.09... \rceil = 5$, by pigeonhole principle, there exists at least 1 line with at least 5 segments covered. This line must be covered by at least 5 dominos. □

Example 7
Given a set S of n integers, show that there exists a non-empty subset of S such that the sum of all the elements of S is divisible by n.

Thought process

Here, the number of subsets is $2^n - 1$, which is a lot more than needed for pigeonhole principle. But just because two subsets differ by a multiple of n does not allow us to construct a subset with sum divisible by n, so the key is to pick suitable subsets to examine.

Solution

Let $S = \{s_1, s_2, ..., s_n\}$ and consider the $n+1$ subsets

$$T_0 = \varnothing, T_1 = \{s_1\}, T_2 = \{s_1, s_2\}, \ ..., T_{n-1} = \{s_1, s_2, ..., s_{n-1}\}, T_n = S$$

Define also a_i to be the sum of all the elements of subset T_i modulo n. By pigeonhole principle, there exists p, q, $0 \le p < q \le n$ such that $a_p \equiv a_q \pmod{n}$. But this means

$$\sum_{i=1}^{p} s_i \equiv \sum_{i=1}^{q} s_i \pmod{n}$$

and therefore

$$\sum_{i=p+1}^{q} s_i \equiv 0 \pmod{n}$$

so our subset is $T_q \setminus T_p = \{s_{p+1}, s_{p+2}, ..., s_q\}$. □

It is worth looking at the difference in flavour between the previous two examples.

Example 6 is heavy on "counting" and relies on careful use of the precise numbers in the question to draw the desired conclusion. This is more common in Combinatorics-type questions, and these also use the strong form of pigeonhole principle more frequently.

Example 7 focuses more on "existence" of some property or some quantity with a certain property, and may use this property to draw further conclusions. This is more common in e.g. Modular Arithmetic-type questions, and sometimes these questions may even have infinite pigeons. The difficulty frequently lies in careful selection of pigeons and/or the steps after applying pigeonhole principle: *okay, we have shown that there are two things with a certain property, now what can I do with them*?

Thus far, H3 A-level exam questions have tended to be more like the latter.

Finally, try writing out an answer for Example 8 first, before reading Example 9.

Example 8
Prove that at a gathering of any 6 people (where any two are either mutual acquaintances or do not know each other), some three of them are either mutual acquaintances or complete strangers to one another.

Example 9
6 points are in general position in space such that no three forms a line. The 15 line segments joining them in pairs are drawn, and then painted with some segments blue and the rest red. Prove that there is (at least) one triangle with all its sides the same colour.

Thought process

Sometimes, when the numbers involved are small, it might be faster to name what you need.

Solution

Let the points be A, B, C, D, E, F. Consider point A. By pigeonhole principle, among the 5 line segments incident from A – AB, AC, AD, AE and AF – at least 3 have the same colour. WLOG, assume that AB, AC and AD are blue. Then if any of BC, BD or CD are blue, then we will have formed a blue triangle. If none of them are blue, then BCD is a red triangle. □

Question to ponder

Clearly Examples 8 and 9 are equivalent problems, but how did the context affect the way we write our solution?

Tutorial Questions

1 Each cell in a 4 by 4 square is filled with one of the numbers 1, 2 or 3. Prove that of the sums along the rows, the columns and the diagonals, two sums must be equal.

Can we achieve this if we consider the sums along the rows and the columns only? That is, if we consider only the sums along the row and columns only, will we always find two sums that are equal?

2 Given 16 distinct positive integers, each less than 62, show that at least three pairs of them have the same absolute difference (the pairs are distinct but need not be disjoint as sets, for e.g., {2, 3} and {3, 4} are considered as two pairs with the same absolute difference).

3 Prove that if 52 distinct integers are chosen, there are two of them whose sum or difference is divisible by 100. (Hint: Consider first the related problem of showing that if 52 distinct integers are chosen from {0, 1, 2, …, 99}, there are two of them whose sum equals 100.)

4 Given any 10 numbers from the set {1, 2, 3, ..., 100}, show that we can always pick two disjoint sets of numbers from the 10 numbers chosen such that they have the same sum. For example, if the 10 numbers chosen are {7, 11, 21, 22, 26, 49, 69, 72, 88, 98}, then we can find two sets {21, 22, 26} and {69} with the same sum.

5 What is the largest subset of {1, 2, 3, …, 2*n*} that can be chosen, so that no member of the subset is a multiple of another member? Justify your answer.

6 Let $S = \{1, 2, \ldots, 2022\}$. Find the smallest possible *k*, such that any subset *A* with $|A| = k$ contains two numbers *a* and *b* with the property $a = 3^s b$ for some integer *s*.

7 [RI 2019 Q6(part)] A lattice point is a point such that its coordinates are all integers.

(i) Prove that among any set of five lattice points (x_1, y_1), (x_2, y_2), …, (x_5, y_5) in the plane, there must be two of them whose midpoint is also a lattice point.

(ii) Determine with proof, the least *m* such that among any set of *m* lattice points (x_1, y_1, z_1), (x_2, y_2, z_2), …, (x_m, y_m, z_m) in space, there must be two of them whose midpoint is also a lattice point.

(iii) The centroid of 4 lattice points (x_a, y_a), (x_b, y_b), (x_c, y_c), (x_d, y_d) is defined as

$$\left(\frac{x_a + x_b + x_c + x_d}{4}, \frac{y_a + y_b + y_c + y_d}{4}\right).$$

Using (i) or otherwise, prove that among any set of 13 lattice points in the plane, there must be four of them whose centroid is also a lattice point.

8 Find the largest integer *n* satisfying the following condition: there exists *n* positive integers $a_1, a_2, \ldots, a_n$ such that $|a_i - a_j|$ is a product of two distinct prime integers for all $i \neq j$.

9 [DHS-EJC 2021 Q7(c)] Let p be a prime number larger than 7. It is a fact that infinitely many terms of the sequence

$$7, 77, 777, 7777, \ldots$$

are divisible by p. One way to prove this is to show that the following two statements are both true:

Statement A: At least one term of the sequence is divisible by p.

Statement B: If there exists one term of the sequence that is divisible by p, then there are infinitely many terms that are divisible by p.

Prove the two statements.

10 John eats at least one apple a day. If he does that for 26 consecutive days, and the total number of apples that he eats is 45, show that there is a sequence of consecutive days that he eats 26 apples altogether.

11 Prove the Erdös-Szekeres Theorem: Given any positive integers r and s, any sequence of distinct real numbers with length at least $rs+1$, contains either an increasing subsequence of length $r+1$ or a decreasing subsequence of length $s+1$ (or both).

(A subsequence is obtained by selecting and keeping some of the terms of the sequence, preserving the order in which these terms appeared in the original sequence. The terms need not be consecutive in the original sequence. For instance, the longest increasing subsequence contained in the sequence $1, 5, 3, 7, 4, 6, 2$ *is* $1, 3, 4, 6$*.)*

Suggested Solutions (Pigeonhole Principle Tutorial)

<table>
<tr><td>1</td><td>
There are altogether 10 sums (4 rows, 4 columns, 2 diagonals). The smallest possible sum is 4 and the largest possible sum is 12, so the number of possible values that the sum can take is $12-4+1=9$.

By pigeonhole principle, since $10>9$, there must be at least 2 sums that are equal in value.

The following example shows that considering sums only along rows and columns is insufficient.

<table>
<tr><td>1</td><td>1</td><td>1</td><td>2</td></tr>
<tr><td>1</td><td>1</td><td>2</td><td>2</td></tr>
<tr><td>2</td><td>3</td><td>3</td><td>3</td></tr>
<tr><td>3</td><td>3</td><td>3</td><td>3</td></tr>
</table>
</td></tr>
<tr><td>2</td><td>
Given 16 distinct positive integers, there are $\frac{16\times 15}{2}=120$ pairs.
The absolute difference of a pair is an integer value between 1 and 60.

Since a difference of 60 can only arise from 1 pair $\{1, 60\}$, there are at least 119 pairs with an absolute difference between 1 and 59.
By pigeonhole principle, there must be $\left\lceil \frac{119}{59} \right\rceil = 3$ pairs with the same absolute difference.
</td></tr>
<tr><td>3</td><td>
Let the set of the 52 integers be I. Construct the following 51 subsets:
$S_0=\{x\in I : x\equiv 0 \ (\mathrm{mod}\,100)\}$
$S_1=\{x\in I : x\equiv 1 \ or \ x\equiv 99 \ (\mathrm{mod}\,100)\}$
$S_2=\{x\in I : x\equiv 2 \ or \ x\equiv 98 \ (\mathrm{mod}\,100)\}$
$S_3=\{x\in I : x\equiv 3 \ or \ x\equiv 97 \ (\mathrm{mod}\,100)\}$
$S_4=\{x\in I : x\equiv 4 \ or \ x\equiv 96 \ (\mathrm{mod}\,100)\}$
…
$S_{49}=\{x\in I : x\equiv 49 \ or \ x\equiv 51 \ (\mathrm{mod}\,100)\}$
$S_{50}=\{x\in I : x\equiv 50 \ (\mathrm{mod}\,100)\}$
By pigeonhole principle, there must be at least 2 elements of I that are in the same subset.
Let x and y be 2 elements of I in the same subset. There are two cases:
<u>Case 1</u>: $x\equiv y \ (\mathrm{mod}\,100)$
$x-y$ is a multiple of 100.
<u>Case 2</u>: $x\not\equiv y \ (\mathrm{mod}\,100)$
By construction, we must have $y\equiv 100-x \ (\mathrm{mod}\,100)$. This means that $x+y\equiv 0 \ (\mathrm{mod}\,100)$, so $x+y$ is a multiple of 100.

Thus we have found two of the integers with the required property.
</td></tr>
</table>

4	There are altogether $2^{10} = 1024$ subsets of the 10 integers chosen, but there can only be 956 possible sums, namely 0, 1, 2, ... , 955. ($955 = 91 + 92 + \ldots + 100$) So by pigeonhole principle, there are two subsets with the same sum. [Alternatively Consider the $2^{10} – 2 = 1022$ subsets with between 1 and 9 elements (inclusive), where there are only 864 possible sums. ($864 = 92 + 93 + \ldots + 100$) So by pigeonhole principle, there are two subsets with the same sum.] Call these two subsets with equal sums X and Y (note that they might not be disjoint). Let $W = X \cap Y$. Then $X \setminus W$ and $Y \setminus W$ are 2 disjoint subsets with the same sum.
5	Each of the $2n$ numbers can be represented in the form $2^k \times m$, where k is a non-negative integer and m is an odd number. In these representations, since m cannot exceed $2n$, there are exactly n distinct values that m can take. Suppose we have selected a subset of $n+1$ or more members. By pigeonhole principle, there must be two members which share the same m when written out in the above representation. But then the larger one will be equal to the smaller one multiplied by a power of 2, so will be a multiple of the smaller one. Thus, we cannot choose a subset with $n+1$ or more members. It remains to be shown that there exists a subset of n members such that no member is a multiple of another member. By inspection, $\{n+1, n+2, ..., 2n\}$ meets this condition.
6	We call n a “three-free” number if the prime factorization of n does not contain 3. There are exactly $2022 - \left\lfloor \frac{2022}{3} \right\rfloor = 1348$ “three-free” numbers in $\{1, 2, \ldots, 2022\}$. For each of these 1348 numbers n, define $T_n = \left\{k \in \mathbb{Z}^+ : k = 3^x n, k \le 2022, x \in \mathbb{Z}^+\right\}$ Then the property $b = 3^s a$ holds if and only if a and b belong in the same T_n, and the 1348 T_n ’s form a partition of S. We can see that $k = 1348$ is not enough as we can simply take the “three-free” numbers. If $k = 1349$, then by pigeonhole principle, at least one of the T_n must have at least 2 numbers, and thus the property holds. Thus, the smallest k is 1349.

7	(i) Consider the parity of each coordinate, which is either odd or even. Hence there are only four pigeonholes (odd, odd), (odd, even), (even, odd) or (even, even). Since there are five lattice points, there must be two of them whose parity for each coordinate are the same. The sum of the two coordinates must thus be even and thus their midpoint $\left(\frac{x_1+x_2}{2},\frac{y_1+y_2}{2}\right)$ is an integer. (ii) Similarly, least $m=2^3+1=9$. Since there are 8 pigeonholes (2 choices for each coordinate), if we have 9 lattice points there must be two of them points whose parity for each coordinate is the same. The midpoint is an integer using the same argument in (i). We will show 8 points is insufficient by a counter-example. Consider the 8 vertices of the unit cube. It is clear that the midpoint of each edge is not a lattice point. In fact, along the edge which the *x*, *y* or *z* coordinate increases by 1, the midpoint along that coordinate is not an integer. (iii) Using the result in (i) to any five of the 13 lattice points we have a pair of lattice points such that their midpoint is a lattice point. Call these two points P_1,P_2. Now with 11 remaining lattice points, apply the result again to get another two lattice points P_3,P_4 whose midpoint is also a lattice point. Continue in this manner until we have 5 remaining lattice points where we apply the result one more time to get another 2 lattice points P_9,P_{10}. Now consider the midpoints of these 5 pairs of lattice points. Call them M_1,M_2,M_3,M_4,M_5. They are all lattice points. Now applying the result in (i) one last time to these 5 points, we can find two of them whose midpoint is a lattice point. However, the midpoint of these two lattice points is the centroid of the original 4 points which the 2 pairs came from. WLOG, if M_1,M_2 are the two lattice points, then we already know they are of the form $M_1\left(\frac{x_a+x_b}{2},\frac{y_a+y_b}{2}\right)$ and $M_2\left(\frac{x_c+x_d}{2},\frac{y_c+y_d}{2}\right)$. Then the midpoint of M_1,M_2 is $$\left(\frac{\frac{x_a+x_b}{2}+\frac{x_c+x_d}{2}}{2},\frac{\frac{y_a+y_b}{2}+\frac{y_c+y_d}{2}}{2}\right)=\left(\frac{x_a+x_b+x_c+x_d}{4},\frac{y_a+y_b+y_c+y_d}{4}\right),$$ which is the centroid of the original 4 points.

8	(Solution by Hannah Tan 25-E4) The difference between any two of the a_n's cannot be a multiple of 4, since no multiple of 4 is a product of two distinct primes. Thus the a_n's are all distinct modulo 4, which means that $n \le 4$. To show $n = 4$, we consider $(a_1, a_2, a_3, a_4) = (11, 33, 66, 88)$ where the differences are $22 = 2\times 11, 33 = 3\times 11, 55 = 5\times 11, 77 = 7\times 11$ which fulfils all the conditions. (Original solution, aka why this appeared in a pigeonhole principle tutorial) If $n \ge 5$, then by pigeonhole principle there must be at least 3 of the a_n's with the same parity (i.e. either all odd or all even). Assume without loss of generality that these are $a_1 < a_2 < a_3$. Then $a_2 - a_1 = 2p_1$, $a_3 - a_2 = 2p_2$ for two odd primes p_1 and p_2, but then we get $a_3 - a_1 = 2p_2 + 2p_1 = 2(p_1 + p_2)$ which is not a product of two distinct primes (since $p_1 + p_2$ is even). Contradiction, so $n < 5$. To show $n = 4$, … <continue as in proof above>
9	**Statement A**: Consider the first $p+1$ terms of the sequence modulo *p*. By pigeonhole principle, at least two of these terms must be congruent to each other modulo *p*. In other words, there are two positive integers *a*, *b* satisfying $0 < a < b \le p+1$ such that $$\overbrace{77...7}^{a\text{ digits}} \equiv \overbrace{77...7}^{b\text{ digits}} \pmod p$$ Taking the difference, we obtain $$\overbrace{77...7}^{b-a}\overbrace{00...0}^{a} \equiv 0 \pmod p$$ However, since $\overbrace{77...7}^{b-a}\overbrace{00...0}^{a} = \overbrace{77...7}^{b-a}\times 10^a = \overbrace{77...7}^{b-a}\times 2^a \times 5^a$, and since *p* is relatively prime to both 2 and 5, we must have $$\overbrace{77...7}^{b-a\text{ digits}} \equiv 0 \pmod p.$$ i.e. the $(b-a)^{\text{th}}$ term is divisible by *p*. **Statement B**: Suppose the c^{th} term is divisible by *p*. Consider the c^{th}, $2c^{\text{th}}$, $3c^{\text{th}}$, … terms of the sequence. Since these are of the form $$\overbrace{77...7}^{c\text{ digits}}\overbrace{77...7}^{c\text{ digits}}\cdots\cdots\overbrace{77...7}^{c\text{ digits}},$$ they are all multiples of $\overbrace{77...7}^{c\text{ digits}}$ and hence divisible by *p*. Thus, there are infinitely many terms in the sequence which are divisible by *p*.

10	Let a_i be the number of apples eaten from the 1st day up to the ith day, where $i = 1, 2, ..., 26$. Additionally, define $a_0 = 0$. By pigeonhole principle, at least two of the a_i, say a_m, a_n where $a_m < a_n$, must be congruent modulo 26. Since $a_n \le 45$, we cannot have $a_n - a_m \ge 52$, so we know that $a_n - a_m = 26$ exactly. But this means that the number of apples eaten from day $m+1$ till day n inclusive is exactly 26. (shown)
11	(Adapted from solution by He Jizhao 21-U5) Firstly, it suffices to show the result for sequences of length exactly $rs+1$, since we can simply take the first $rs+1$ terms of any longer sequence. Next, the result we have to show is equivalent to the statement "Any sequence of distinct real numbers with length $rs+1$, which does not have an increasing subsequence of length $r+1$, must have a decreasing subsequence of length $s+1$". We will prove this version. Denote the terms of the sequence by $a_1, a_2, ..., a_{rs+1}$. Let l_i, $i = 1, 2, ..., rs+1$ denote the length of the longest increasing sequence starting from a_i. Since there are no increasing sequences of length $r+1$, then $1 \le l_i \le r$ for all *i*. By pigeonhole principle, there must be $\left\lceil \frac{rs+1}{r} \right\rceil = s+1$ distinct terms of these satisfying $l_{n_1} = l_{n_2} = ... = l_{n_{s+1}} = l$. Consider the terms a_{n_1} and a_{n_2}. We cannot have $a_{n_1} < a_{n_2}$ since there is an increasing sequence of length *l* with first term a_{n_2}; by adding a_{n_1} to the front we will obtain an increasing sequence of length $l+1$ with first term a_{n_1}, which contradicts the fact that the longest increasing sequence with first term a_{n_1} has length *l*. Therefore, we must have $a_{n_1} > a_{n_2}$ since all the terms are distinct. Using the same argument, we must also have $a_{n_2} > a_{n_3}$, and indeed we have $a_{n_1} > a_{n_2} > a_{n_3} > ... > a_{n_{s+1}}$ which is a decreasing sequence of length $s+1$. (shown) (Alternative solution) Denote the terms of the sequence by $a_1, a_2, ..., a_{rs+1}$. Let x_i, $i = 1, 2, ..., rs+1$ denote the length of the longest increasing sequence starting from a_i, and y_i denote the length of the longest decreasing sequence starting from a_i. We claim that it is not possible to have $(x_i, y_i) = (x_j, y_j)$ for any $i < j$. Proof: <u>Case 1</u>, $a_i < a_j$: We can add a_i to the front of the longest increasing sequence starting at a_j, so the longest increasing sequence starting at a_i has to be longer than the longest increasing sequence starting at a_j. In other words, $x_i > x_j$.

Case 2, $a_i > a_j$: Similarly, we can add a_i to the front of the longest decreasing sequence starting at a_j, so $y_i > y_j$.
In both cases, $(x_i, y_i) \neq (x_j, y_j)$. [Claim proven]

Suppose for a contradiction that $x_i \leq r$ and $y_i \leq s$ for all *i*. Then there are only rs different combinations for (x_i, y_i), so by pigeonhole principle we must have $(x_i, y_i) = (x_j, y_j)$ for some *i* and *j* which violates the claim above. Contradiction.

So we must have either $x_i \geq r+1$ or $y_i \geq s+1$ for some *i*, i.e. there must be either an increasing sequence of length $r+1$ or a decreasing sequence of length $s+1$.
(shown)